Vocal markers of motor, cognitive, and depressive symptoms in Parkinson's disease

Kara M. Smith
Department of Neurology
University of Massachusetts Medical School
Worcester, MA, USA
Kara.Smith@umassmemorial.org

James R. Williamson and Thomas F. Quatieri
Bioengineering Systems and Technologies
MIT Lincoln Laboratory
Lexington, MA, USA
[jrw,quatieri]@ll.mit.edu

***Abstract*—In addition to motor dysfunction, Parkinson's disease (PD) often results in symptoms of cognitive impairment and depression, which can go underdiagnosed and undertreated. One approach that may improve diagnosis and differentiation of motor, cognitive, and depressive symptoms of PD relies on vocal acoustics that has previously been used to predict symptoms in each of these domains separately. In this paper, a joint multi-domain characterization of the PD symptoms is presented. Speech recordings from 35 PD patients were analyzed for speech markers characterizing articulatory coordination based on resonant (formant) frequencies and delta-mel cepstral coefficients (dMFCC), as well as phonemic timing based on phoneme-dependent speaking rates. Moderate correlations were found between vocal markers and the motor and cognitive symptoms of PD, and weaker correlations with depressive symptoms. We identified notable differences in the correlation patterns, suggesting it may be possible to distinguish the impact of different PD symptoms on speech. Statistical models based on the vocal markers achieved moderate accuracy in predicting motor severity (r=0.42) and global cognition (r=0.52) but not depression (r=-0.21). Future study is warranted to further develop symptom-specific vocal marker models in PD.**

1. Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disease and affects one million Americans. Although PD is most often characterized by its motor symptoms, non-motor symptoms are prevalent and highly impactful. Common non-motor symptoms include depression, which occurs in 35% of patients with PD [1] and cognitive impairment, which occurs in 80% of patients withPD over their disease course [2]. Depression and cognitive impairment are associated with more severe disease course, impaired quality of life, and increased mortality [3-5] and may interact with the motor features of PD. Depression causes psychomotor slowing and flat affect that mimics the characteristic bradykinesia and facial masking that occurs in PD. Patients with PD often have a monotonous voice with diminished prosody, which shares features with voice changes in depression [6, 7]. Likewise, cognitive changes in PD often involve executive dysfunction, which can impact verbal fluency [9], as well as gait and balance [10]. As a result, current clinical assessment tools are unable to clearly distinguish the impact of these non-motor symptoms and motor symptoms on global function and thus are underdiagnosed and undertreated in PD [8]. Clinicians need objective and precise tools to assess neuropsychiatric non-motor symptoms in PD. Vocal markers could be a powerful tool to meet these challenges and help to disentangle the underlying neurophysiology of mood, cognitive, and motor impairment in PD.

Although there has been growing interest in using automated voice analysis to detect and monitor PD disease status [11-18], this body of work has focused largely on motor impairments that are the most commonly recognized changes in speech in PD. These include imprecise articulation, monotonous and reduced pitch and volume, variable speech rate and pause segments, breathy and harsh voice quality, and changes in intonation and rhythm [11]. Almost all patients with PD experience motor vocal changes over their disease course. Vocal markers based on these changes can distinguish PD from healthy controls [12-14] and predict PD disease severity [14-18]. However, this work, in focusing on motor symptoms, has not accounted for affective and cognitive influences on speech. To develop useful vocal markers in PD, the motor, cognitive and affective components of speech need to be better distinguished and understood.

Motor and neuropsychiatric deficits in PD are thought to reflect different underlying neurochemical and

This material is based upon work supported by the Assistant Secretary of Defense for Research and Engineering under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Assistant Secretary of Defense for Research and Engineering.

neuropathological etiologies. Motor symptoms are due to degeneration of dopaminergic neurons in the substantia nigra, with resulting dopaminergic deficit in the basal ganglia. Depression in PD is a more heterogeneous process, involving dopamine [19], serotonin [20], and norepineprhine [21] depletion, and structural changes in limbic regions [22, 23]. Cognitive impairment is related in part to a dopamine deficit, which causes dysfunction of the cognitive pathways connecting the frontal lobes and basal ganglia. This impacts executive function, notably including lexical retrieval and processing speed [24, 25]. However, many other factors play a role in cognition in PD including other neurotransmitter levels, and both PD-related Lewy pathology and amyloid protein deposition in the cortex [26]. Additionally, functional neuroimaging has revealed distinct patterns of resting state metabolism associated with cognitive and depressive symptoms in PD [27-29]. We therefore sought to identify vocal markers of depression and cognition in PD reflective of these differences in pathophysiology.

We have previously identified vocal markers of depression in the general population [30-32], utilizing changes in phonation, articulation and prosody known to occur across different time scales and speech segments to reflect the underlying neurophysiology of speech production. In this previous work, we used the correlation structure of formant frequencies and delta-mel frequency cepstral coefficients (dMFCCs) to represent underlying changes in vocal tract shape and dynamics, as well as phoneme-dependent speaking rates, to predict symptom severity in major depressive disorder. Using a similar approach, we have also previously identified vocal markers associated with cognitive impairment, verbal fluency, and cognitive load in the general population [33-35]. In PD, linguistic changes associated with cognitive impairment include decreased speech rate, increased pausing between utterances, and impaired grammar, but vocal acoustics related to these symptoms have not been explored with a data-driven approach [36, 37].

In the current work, we apply our timing- and coordination-based vocal acoustic feature sets and data-driven analytical approach in patients with PD to predict motor, cognitive, and depressive symptoms in PD. We hypothesize that these PD symptoms will correlate with specific, identifiable changes in vocal tract shape and dynamics and are dependent on articulatory and phonetic categories. Additionally, we aim to explore the differential impact of motor impairment, cognitive impairment, and depression on PD speech. We hypothesize that while some features would overlap across these domains, others will be distinctly correlated with each domain. Our long-term goal is to establish symptom-specific feature clusters that will fuel the growing field of vocal biomarkers in PD.

2. Methods

2.1. Audio Recordings

We studied 35 PD patients, who were enrolled at the Perelman School of Medicine at the University of Pennsylvania. PD was diagnosed according to published criteria [38]. All subjects completed an informed consent procedure in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the University of Pennsylvania.

Motor deficits were characterized by the Unified Parkinson's disease Rating Scale (UPDRS) obtained within 6 months of speech testing. Global cognition was assessed with the Montreal Cognitive Assessment (MoCA). Patients with dementia were excluded based on MoCA score <24 [39]. Depressive symptoms were assessed with the Geriatric Depression scale (GDS) [40]. Table I summarizes basic assessment statistics from the 35 patients, and Table II indicates between-outcome Spearman correlations.

TABLE I. Statistics on assessments used to measure PD symptoms in motor (UPDRS), cognitive (MoCA), and depression (GDS) domains.

Assessment	**Min**	**Max**	**Mean**	**Std. Dev.**
UPDRS	3	53	20.46	10.45
MoCA	24	30	27.31	2.03
GDS	0	13	2.46	2.94

TABLE II. Spearman correlations between assessment outcome measures.

Assessments	**r**	**p**
UPDRS & MoCA	-0.38	0.02
UPDRS & GDS	0.27	0.12
MoCA & GDS	0.01	0.94

2.2. Feature Extraction

We utilize three feature sets, described below, which have been shown in our previous research to be predictive of PD motor severity [17], major depressive disorder [30-32], and cognitive performance [33, 34]. These feature sets characterize articulatory coordination based on the dynamics of vocal resonant frequencies and spectral properties, as well as changes in vocal timing at the phoneme level.

Articulatory Coordination

Formant frequency tracks. Properties of vocal tract resonances over time contain information about speech dynamics related to articulatory properties of the depressed voice. A formant tracking algorithm based on Kalman

filtering was used to obtain smooth estimates of the first three resonant frequencies over time [42]. Formant frequencies were extracted every 10 ms from the audio signal. Embedded in the formant tracking algorithm is a voice-activity detector that allows a Kalman smoother to smoothly coast through non-speech regions. Estimates of the third formant that went above a threshold of 4.5k Hz were truncated.

Articulatory coordination was estimated from formant tracks using correlation structure features. In this feature approach, a channel-delay correlation matrix is computed from the formant tracks using time-delay embedding. The correlation matrix has dimensionality (45 x 45), based on three formant channels and 15 time delays per channel. A single delay scale with 7-frames (70 ms) delay spacing is used. From the correlation matrix a 45-dimensional rank ordered eigenspectrum is computed, which characterizes the within-channel and cross-channel distributional properties of the multivariate formant time series.

Delta Mel Frequency Cepstral Coefficients. To introduce vocal tract spectral magnitude information, a standard set of 16 Mel Frequency Cepstral Coefficients (MFCCs) was generated by Opensmile from segmented but otherwise unprocessed audio files [43]. Delta MFCCs (dMFCCs) were then computed, which reflect dynamic velocities of the MFCCs over time. Delta coefficients were computed using regression over a 5-frame window. A channel-delay correlation matrix was computed from the dMFCCs using time-delay embedding, with dimensionality (240 x 240), based on 16 dMFCC channels and 15 delays per channel with 1-frame (10 ms) delay spacing. From this matrix the 240-dimensional rank-ordered eigenspectrum was computed, which characterizes the within-channel and cross-channel distributional properties of the multivariate dMFCC time series.

Phonetic Timing: We have found that computing phoneme specific characteristics, rather than average measures of speaking rate, reveals stronger relationships between speech rate and depression severity [31, 45]. Using an automatic phoneme recognition algorithm [44], we detect phonetic boundaries and phoneme specific durations that are associated with each instance of the 40 classes of defined phonetic speech units. Consistent with previous work on free speech recordings that have variable total durations, the summed durations of each phoneme are normalized by the total number of phonemes to produce a phoneme rate measure.

Based on previous work [31, 45], for the phonemes that have rates that are highly correlated with an outcome measure (UPDRS, MoCA, or GDS) on the training set, the rates are linearly combined to yield a fused phoneme rate measure, with the sign of the combination weight based on the correlation sign. Consistent with [45], we use binary weights of 1 or -1 to combine the rates of the selected phonemes. In the current work, we select the top eight correlating phonemes, since eight is the average of the number of phonemes (six and ten) use on two different recorded passages in [31].

2.3. Correlational Analysis

The vocal markers consist of high dimensional feature vectors: 45-dimensional formant eigenspectra, 240-dimensional dMFCC eigenspectra, and 40-dimensional phoneme rate measures. As a first level of analysis, the correlations of the raw feature elements with the three PD symptom outcomes are measured. Specifically, correlations are measured with motor symptoms (UPDRS scores), cognitive impairment symptoms (MoCA scores), and depression symptoms (GDS scores). Because MoCA scores are negative indicators of impairment, correlatons are done with negative MoCA scores, for sign consistency with the UPDRS and GDS correlations.

2.4. Outcome Prediction

Regression Model. A standard statistical approach called Gaussian staircase (GS) regression [17, 30, 31] is used for prediction. GS generalizes the use of a Gaussian classifier for regression into an ensemble of Gaussians for each class (Class 1 = “lower”, Class 2 = “higher”), based on partitioning of the range of values of the outcome variable. The ensemble of Gaussians associated with each class is interpreted as a Gaussian mixture model, such that the Class 1 or Class 2 likelihood for a test data point is the sum of likelihoods from the Gaussian ensemble associated with that Class. The outcome prediction is then obtained using univariate regression based on the two-class log-likelihood scores, with the regression model constructed from the training set log-likelihood scores.

For this work, GS levels were determined for each outcome variable as follows. The UPDRS Class 1 partitions were [0-11, 0-16, 0-21, 0-26], the MoCA Class 1 partitions were [0-25, 0-26, 0-27, 0-28, 0-29], and the GDS Class 1 partitions were [0-0, 0-1, 0-2, 0-3, 0-4]. For each symptom domain, the Class 2 partitions are simply the complement of the Class 1 partitions. Regularization was done by adding covariance values of 10 to the diagonal elements of the Gaussian covariance matrices. 2^{nd}-order univariate regression was used to map log-likelihoods to outcome predictions.

Dimensionality Reduction. In order to avoid possible feature selection biases on a relatively small (35-subject) dataset, we adopted the same feature selection parameters that were used in the depression prediction system that won first place in the AVEC 2014 depression prediction sub-challenge [31]. Specifically, the first four principal component features were used for the formant correlation

structure features, and the first five principal components for the delta-MFCC correlation structure features.

In the AVEC 2014 system, phoneme rate features were used on two different passages, a read passage and a free speech passage. The top six correlating phonemes were combined in the read passage, and the top ten phonemes in the free speech passage. Here, we split the difference, using the top eight correlating phonemes.

Cross-validation. In order to obtain an unbiased estimate of generalization performance, a cross-validation procedure is used such that statistical models are trained on rotating data subsets, and applied to held-out test data, with no overlap in subject identity between training and test sets. Within this procedure, all transformations that depend on features obtained across multiple recordings, such as z-scoring, PCA, and correlation-based phoneme rate aggregation, are computed strictly within the training set and then applied to the held-out test set.

One shortcoming of small data sets is that randomly partitioned cross-validation folds, or leave-one-subject cross-validation, can result in negatively biased estimates, since the training set outcome variables tend to be negatively correlated with the test set outcome variable. To avoid this complication, a 12-fold stratified sampling cross-validation procedure is used, in which the expected value of the outcome variable is kept as consistent as possible across the different test folds.

3. Results

3.1. Correlational Analysis

Figure 1 shows the Spearman correlations between the formant eigenvalue features (top) and the dMFCC eigenvalue features (bottom) with the three outcomes. The eigenvalues are rank ordered left to right from largest to smallest. Overall, smaller formant eigenvalues are negatively correlated with symptom severity, and smaller dMFCC eigenvalues are positively correlated with symptom severity. This correlation pattern is stronger for motor and cognitive symptoms, and weaker for depressive symptoms.

This pattern of results is consistent with previous work showing similar patterns in formant- and dMFCC-based eigenvalue features for motor symptoms in Parkinson's disease [17] and for symptoms of major depressive disorder [30, 31]. Similar results have also been found in formant-based eigenavalue features for symptoms of reduced cognitive performance related to aging [34] and possible mTBI [33].

Despite the overall similarities in correlation patterns among the various PD symptom domains, there are some notable differences. First, the depression symptoms show smaller absolute correlation levels. This may be due to the low and restricted range of GDS scores (see Table 1), indicating minimal depressive symptoms are present in this cohort. Second, there are notable shifts in the UPDRS versus MoCA correlation patterns. For the formant-based eigenvalues, there are stronger negative correlations with UPDRS among eigenvalues 5 through 20. For the dMFCC-based eigenvalues, on the other hand, there are stronger correlations with MoCA among eigenvalues 20 through 40, and stronger correlations with UPDRS among eigenvalues 54-140 and 220-240. These differences indicate that there may be subtle differences in the effects of motor and cognitive symptoms on articulatory speech dynamics.

Figure 2 shows the Spearman correlations between MoCA outcomes. For clarity, correlations with GDS are not included, as these correlations were considerably smaller. The phonemes are divided into three categories: 1) phonemes with similar outcome correlations (left); 2) strong UPDRS and weak MoCA correlations (center); 3) weak UPDRS and strong MoCA correlations (right). In the first category are phonemes indicating that motor and cognitive symptoms are positively associated with: slower speech planning and execution ('sil', 'ah'), slower labial, labial/dental, and/or tongue/dental movements ('uw', 'th', 'b', 'v'), and faster

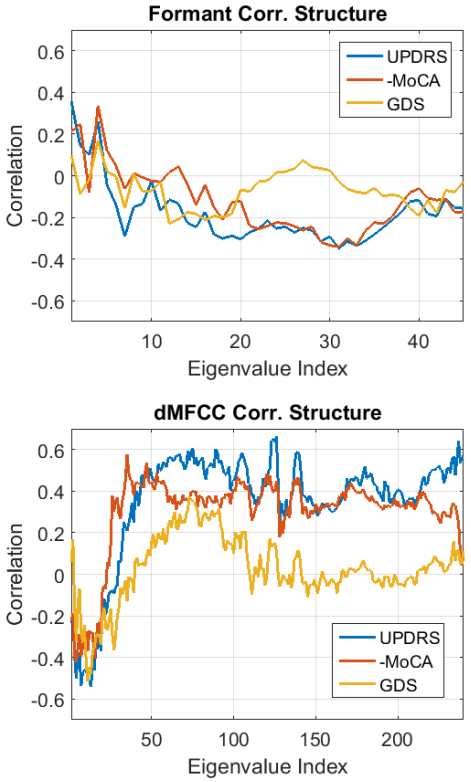

Figure 1. Spearman correlations of formant-based (top) and dMFCC-based (bottom) eigenvalue features with three PD symptom outcomes. Eigenvalues are ordered, largest to smallest, from left to right, and MoCA is sign-adjusted.

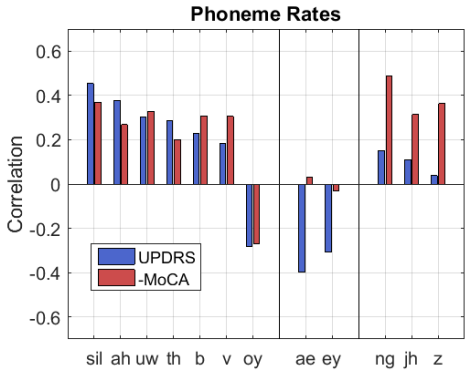

Figure 2. Spearman correlations of phoneme rates with motor and cognitive impairment symptoms shown for those phonemes that have |r|>0.28 for either outcome.

execution of the dipthong vowel 'oy'. In the second category are two open-vowel phonemes ('ae', 'ey') that require large jaw movements, for which faster rates are positively associated with motor symptoms only. For the phonemes 'oy', 'ae', and 'ey', the positive association between faster rates and symptom severity could be due to shorter or less complete motor trajectories. Finally, in the third category are three consonants ('ng', 'jh', 'z'), which require precise tongue articulation, for which slower rates are positively associations with cognitive symptoms only.

3.2. Outcome Prediction

Table III shows the correlations obtained in predicting the three outcome variables based on each individual feature set using 12-fold stratified sampling. It also shows the results obtained by combining the three feature sets, which was done by adding the GS log-likelihood ratios across the three feature sets, prior to the univariate regression step. The dMFCC feature set obtained the strongest correlations for both motor and cognitive symptoms. For cognitive symptoms, predictions combining the three feature sets improved prediction performance. None of the feature sets were useful for predicting depression symptoms. Figure 3 illustrates the predicted motor assessments (top) and cognitive assessments (bottom) as a function of true assessment values for the combined system.

TABLE III. Spearman correlations between predicted and true symptom assessment scores, based on vocal markers.

Feature Sets	**UPDRS**		**MoCA**		**GDS**	
	r	***(p)***	***r***	***(p)***	***r***	***(p)***
Formants	0.20	(0.24)	0.16	(0.35)	-0.04	(0.84)
dMFCC	0.43	(0.01)	0.39	(0.02)	0.02	(0.92)
Phonemes	0.13	(0.46)	0.29	(0.09)	-0.45	(0.01)
Combined	0.42	(0.01)	0.52	(0.00)	-0.21	(0.23)

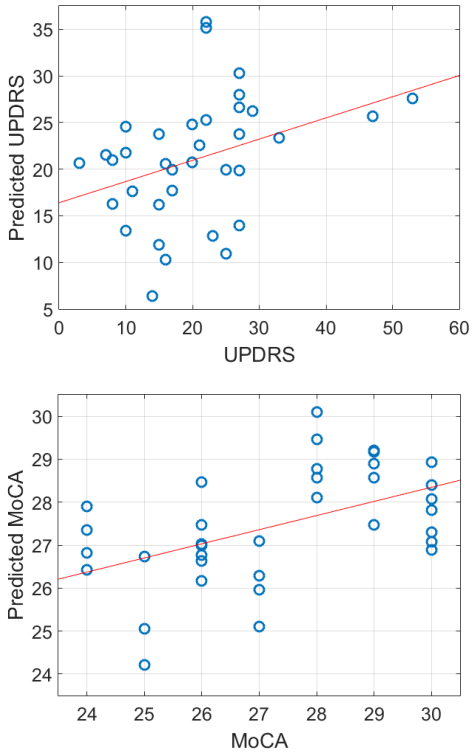

Figure 3. Fused system predictions are plotted as a function of true values for UPDRS (top; r=0.42) and MoCA (bottom; r=0.52).

4. Discussion

Using high-level acoustic features, we identified vocal markers of neuropsychiatric symptoms in PD. Our approach adds to the field of vocal biomarker assessment in several ways. First, we have identified vocal markers associated with depression and cognition in PD for the first time to our knowledge. Second, we established that different characteristic values of formant and dMFCC correlation structure, and of phonemic durations and categories, are correlated with motor and non-motor symptoms. Our results were more robust for cognition compared with depression. Future work is needed to explore the elements of the speech task demands and the optimal features to better assess depression.

Strengths of our approach include the relative independence from patient characteristics such as sex and age, due to the normalization inherent in the correlation feature structure approach. This approach may also more reliably control for variations within individuals with PD. Variability in symptoms is a challenge for PD biomarker development, since PD symptoms may vary based on environment, concurrent motor and/or cognitive load, and medication effects. Another strength is the ecological validity

of our approach. We were able to demonstrate success using an iPhone in typical clinic room or home setting, rather than a lab setting. This suggests that vocal biomarker monitoring may be feasible without expensive or intrusive equipment, allowing it to characterize the patient's unique daily experience. Weaknesses of our study include the limited range of depressive symptom severity in this patient population. Future work should include PD patients with more severe depressive symptoms. In addition, our sample size was relatively small and included only patients recruited at an academic medical center, and may not be generalizable to other PD patient populations.

5. Conclusion

Vocal markers are a promising tool to assess both motor and non-motor, including depressive and cognitive, symptoms in patients with PD. This work supports the feasibility of symptom-specific feature clusters that enhance the further development of vocal biomarkers in PD and suggests correlation dependence on articulatory and phonetic categories. The major advantages of vocal markers are that speech can be tested remotely and automatically, allowing for frequent and quantitative symptom assessment. This approach could fuel large-scale screening of patients and improved monitoring of fluctuating symptoms during daily activities, as well as monitoring of response to therapeutics. Further research is warranted using multi-modality feature analysis with additional motor and affective components in order to better detect neuropsychiatric symptoms in PD.

Acknowledgments

The authors thank Dr. Adam Lammert for helpful discussions about phoneme correlation analysis.

References

[1] Reijnders JS, Ehrt U, Weber WE, Aarsland D, Leentjens AF (2008) A systematic review of prevalence studies of depression in Parkinson's disease. *Mov Disord* **23**, 183-189; quiz 313.

[2] Hely MA, Reid WG, Adena MA, Halliday GM, Morris JG (2008) The Sydney multicenter study of Parkinson's disease: the inevitability of dementia at 20 years. *Mov Disord* **23**, 837-844.

[3] Balestrino R, Martinez-Martin P (2017) Neuropsychiatric symptoms, behavioural disorders, and quality of life in Parkinson's disease. *J Neurol Sci* **373**, 173-178.

[4] Oosterveld LP, Allen JC, Jr., Reinoso G, Seah SH, Tay KY, Au WL, Tan LC (2015) Prognostic factors for early mortality in Parkinson's disease. *Parkinsonism Relat Disord* **21**, 226-230.

[5] Leroi I, McDonald K, Pantula H, Harbishettar V (2012) Cognitive impairment in Parkinson disease: impact on quality of life, disability, and caregiver burden. *J Geriatr Psychiatry Neurol* **25**, 208-214.

[6] Darby JK, Simmons N, Berger PA (1984) Speech and voice parameters of depression: a pilot study. *J Commun Disord* **17**, 75-85.

[7] Flint AJ, Black SE, Campbell-Taylor I, Gailey GF, Levinton C (1993) Abnormal speech articulation, psychomotor retardation, and subcortical dysfunction in major depression. *J Psychiatr Res* **27**, 309-319.

[8] Weintraub D, Moberg PJ, Duda JE, Katz IR, Stern MB (2003) Recognition and treatment of depression in Parkinson's disease. *J Geriatr Psychiatry Neurol* **16**, 178-183.

[9] Henry JD, Crawford JR (2004) Verbal fluency deficits in Parkinson's disease: a meta-analysis. *J Int Neuropsychol Soc* **10**, 608-622.

[10] Yogev-Seligmann G, Hausdorff JM, Giladi N (2008) The role of executive function and attention in gait. *Mov Disord* **23**, 329-342; quiz 472.

[11] Goberman AM, Coelho C (2002) Acoustic analysis of parkinsonian speech I: speech characteristics and L-Dopa therapy. *NeuroRehabilitation* **17**, 237-246.

[12] Rusz J, Cmejla R, Ruzickova H, Ruzicka E (2011) Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *J Acoust Soc Am* **129**, 350-367.

[13] Rusz J, Cmejla R, Tykalova T, Ruzickova H, Klempir J, Majerova V, Picmausova J, Roth J, Ruzicka E (2013) Imprecise vowel articulation as a potential early marker of Parkinson's disease: effect of speaking task. *J Acoust Soc Am* **134**, 2171-2181.

[14] Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO (2012) Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE Trans Biomed Eng* **59**, 1264-1271.

[15] Silbergleit AK, LeWitt PA, Peterson EL, Gardner GM (2015) Quantitative Analysis of Voice in Parkinson Disease Compared to Motor Performance: A Pilot Study. *J Parkinsons Dis* **5**, 517-524.

[16] Tsanas A, Little MA, McSharry PE, Ramig LO (2011) Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average Parkinson's disease symptom severity. *J R Soc Interface* **8**, 842-855.

[17] Williamson JR, Quatieri TF, Helfer BS, et al. (2015) Segment-dependent dynamics in predicting Parkinson's disease. . *Sixteenth Annual Confernece of the International Speech Communication Association*

.[18] Orozco-Arroyave, J. R., Arias-Londoño, J. D., Bonilla, J. F. V., Gonzalez-Rátiva, M. C., & Nöth, E. (2014, May). New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease. In *LREC* (pp. 342-347).

[19] Vriend C, Raijmakers P, Veltman DJ, van Dijk KD, van der Werf YD, Foncke EM, Smit JH, Berendse HW, van den Heuvel OA (2014) Depressive symptoms in Parkinson's disease are related to reduced [123I]FP-CIT binding in the caudate nucleus. *J Neurol Neurosurg Psychiatry* **85**, 159-164.

[20] Hesse S, Meyer PM, Strecker K, Barthel H, Wegner F, Oehlwein C, Isaias IU, Schwarz J, Sabri O (2009) Monoamine transporter availability in Parkinson's disease patients with or without depression. *Eur J Nucl Med Mol Imaging* **36**, 428-435.

[21] Remy P, Doder M, Lees A, Turjanski N, Brooks D (2005) Depression in Parkinson's disease: loss of dopamine and noradrenaline innervation in the limbic system. *Brain* **128**, 1314-1322.

[22] Huang P, Xuan M, Gu Q, Yu X, Xu X, Luo W, Zhang M (2015) Abnormal amygdala function in Parkinson's disease patients and its relationship to depression. *J Affect Disord* **183**, 263-268.

[23] van Mierlo TJ, Chung C, Foncke EM, Berendse HW, van den Heuvel OA (2015) Depressive symptoms in Parkinson's disease are related to decreased hippocampus and amygdala volume. *Mov Disord* **30**, 245-252.

[24] Grossman M, Lee C, Morris J, Stern MB, Hurtig HI (2002) Assessing resource demands during sentence processing in Parkinson's disease. *Brain Lang* **80**, 603-616.
[25] Grossman M, Zurif E, Lee C, Prather P, Kalmanson J, Stern MB, Hurtig HI (2002) Information processing speed and sentence comprehension in Parkinson's disease. *Neuropsychology* **16**, 174-181.
[26] Zgaljardic DJ, Foldi NS, Borod JC (2004) Cognitive and behavioral dysfunction in Parkinson's disease: neurochemical and clinicopathological contributions. *J Neural Transm (Vienna)* **111**, 1287-1301.
[27] Borgonovo J, Allende-Castro C, Laliena A, Guerrero N, Silva H, Concha ML (2017) Changes in neural circuitry associated with depression at pre-clinical, pre-motor and early motor phases of Parkinson's disease. *Parkinsonism Relat Disord* **35**, 17-24.
[28] Hu X, Song X, Yuan Y, Li E, Liu J, Liu W, Liu Y (2015) Abnormal functional connectivity of the amygdala is associated with depression in Parkinson's disease. *Mov Disord* **30**, 238-244.
[29] Huang C, Mattis P, Tang C, Perrine K, Carbon M, Eidelberg D (2007) Metabolic brain networks associated with cognitive function in Parkinson's disease. *Neuroimage* **34**, 714-723.
[30] Williamson, J. R., Quatieri, T. F., Helfer, B. S., Horwitz, R., Yu, B., & Mehta, D. D. (2013, October). Vocal biomarkers of depression based on motor incoordination. In *Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge* (pp. 41-48). ACM.
[31] Williamson, J. R., Quatieri, T. F., Helfer, B. S., Ciccarelli, G., & Mehta, D. D. (2014, November). Vocal and facial biomarkers of depression based on motor incoordination and timing. In *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge* (pp. 65-72). ACM.
[32] Williamson, J. R., Godoy, E., Cha, M., Schwarzentruber, A., Khorrami, P., Gwon, Y., ... & Quatieri, T. F. (2016, October). Detecting Depression using Vocal, Facial and Semantic Communication Cues. In *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge* (pp. 11-18). ACM.
[33] Helfer, B. S., Quatieri, T. F., Williamson, J. R., Keyes, L., Evans, B., Greene, W. N., ... & Palmer, J. (2014). Articulatory dynamics and coordination in classifying cognitive change with preclinical mTBI. In *INTERSPEECH* (pp. 485-489).
[34] Yu B, Quatieri, TF, Williamson, JR, & Mundt, JC (2014) Prediction of cognitive performance in an animal fluency task based on rate and articulatory markers. *INTERSPEECH*, 1038-1042.
[35] Quatieri TF, Williamson, J. R., Smalt, C. J., Patel, T., Perricone, J., Mehta, D. D. et al. (2015) Vocal biomarkers to discriminate cognitive load in a working memory task. *INTERSPEECH*, 2684-2688.
[36] Ash S, McMillan C, Gross RG, Cook P, Gunawardena D, Morgan B, Boller A, Siderowf A, Grossman M (2012) Impairments of speech fluency in Lewy body spectrum disorder. *Brain Lang* **120**, 290-302.
[37] Ash S, McMillan C, Gross RG, Cook P, Morgan B, Boller A, Dreyfuss M, Siderowf A, Grossman M (2011) The organization of narrative discourse in Lewy body spectrum disorder. *Brain Lang* **119**, 30-41.
[38] Hughes AJ, Daniel SE, Kilford L, Lees AJ (1992) Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* **55**, 181-184.
[39] Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* **53**, 695-699
[40] Weintraub D, Oehlberg KA, Katz IR, Stern MB (2006) Test characteristics of the 15-item geriatric depression scale and Hamilton depression rating scale in Parkinson disease. *Am J Geriatr Psychiatry* **14**, 169-175.
[41] Kaplan E, Goodglass, H., & Weintraub, S. (1983) *The Boston Diagnostic Aphasia Examination*, Lea & Febiger, Philadelphia, PA.
[42] Mehta DD RD, Wolfe PJ (2012) Kalman-based autoregressive moving average modeling and inference for formant and antiformant tracking. *The J. of the Acoustical Society of America* **132**, 1732-1746.
[43] Eyben F WF, Gross F, Schuller B (2013) in *ACM Multimedia (MM)*, Barcelona, Spain, pp. 835–838.
[44] Shen, W., White, C. M., & Hazen, T. J. (2009). *A comparison of query-by-example methods for spoken term detection*. MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB.
[45] Trevino AC, Quatieri TF, Malyska N (2011) Phonologically-based biomarkers for major depressive disorder. *EURASIP Journal on Advances in Signal Processing* **1**, 1-18.